Formal Foundation of Induced Coherence Weights:
Compression-based Coherence and Operational Information Measures

Benjamin James

February 23, 2026

Abstract

Coherence-weighted information theory extends Shannon's framework by distinguishing transmissible structure that contributes to persistence from structure that is merely unpredictable. My prior work established coherence-weighted entropy and mutual information, along with capacity and selective compression theorems, conditional on a bounded symbol-weight function $w(x) \in [0,1]$. This third essay addresses the status of w as a principled quantity rather than an arbitrary knob. I specify a minimal set of admissibility commitments that any coherence-derived weighting must satisfy, including boundedness, monotonicity, coarse-graining consistency, recursive stability, empirical derivability, and domain-translation invariance. I then provide a concrete derivation pipeline from observed state streams to proxy coherence estimates $\hat{C}$, from $\hat{C}$ to symbol-level relevance $\rho(x)$, and from $\rho(x)$ to induced weights $w(x)$. The pipeline yields falsifiable predictions. Coherence-weighted compression should preserve coherence-bearing structure at equal bitrate under distribution shift; weight rankings should remain stable under benign coarse-graining; coherence declines should lead regime transitions; and constraint-relief interventions should measurably increase coherence proxies relative to purely interpretive changes. Finally, I isolate the residual hard seam, the choice of compression estimator K, and restate the invariance claim as robustness across estimator classes rather than estimator-independence. The result is a self-contained operational program where coherence weights are not chosen, but induced under explicit constraints, enabling empirical convergence and principled falsification.

Keywords: coherence information theory, coherence weights, induced relevance, symbolic compression, selective compression, coherence entropy, coherence mutual information, toy dynamical system, rate-distortion tradeoff, falsifiable predictions

Any attempt to elevate “meaningful information” into an operational theory inherits an immediate suspicion that relevance is a discretionary label, inserted by the theorist at the point where mathematics would otherwise compel neutrality. Classical information theory avoids this suspicion by refusing relevance altogether. It counts bits regardless of whether their transmission sustains an organism, stabilizes an institution, or reinforces a model that can survive contact with its environment. A coherence-weighted extension rejects that neutrality, claiming that some structure matters because it participates in persistence, but that claim becomes operational only if the weighting that encodes “matters” is itself disciplined.

In the coherence-weighted framework, the weight function $w(x) \in [0,1]$ is the hinge. If it is hand-assigned, the theory remains a flexible language for preference. If it is induced, the theory becomes a constraint-sensitive generalization with predictive bite. This essay isolates precisely what it means for w to be induced, and how such induction can proceed without importing exogenous semantics. The goal is not uniqueness in closed form, but convergence under shared commitments. Two researchers, given the same data and the same admissibility constraints, should be driven toward materially similar weightings, up to monotone transformations and estimator noise.

To say that w is “coherence-derived” must mean more than “bounded and convenient.” It must mean that w is constrained by the coherence kernel in the same way that channel capacity is constrained by mutual information. Not immune to modeling choices, but not freely assignable either. The following commitments are intentionally minimal. They do not force a single estimator, but they prohibit the degrees of freedom that would allow coherence weighting to collapse into interpretive preference.

M1: Boundedness and monotonicity

Weights must be bounded, $w(x) \in [0,1]$, and monotone in a coherence-relevance score $\rho(x)$. Boundedness is not cosmetic; it prevents weights from becoming an unacknowledged rescaling of probability or utility. Monotonicity prevents “relevance” from being defined post

hoc. If $\rho(x)$ increases, $w(x)$ must not decrease. This commitment supplies the first discipline. Weights can vary, but only in the direction coherence relevance demands.

M2: Coherence-consistency under coarse-graining

A coherence claim is structurally void if it depends on the analyst's choice of resolution. If micro-symbols are aggregated into macro-symbols, the induced weights on macro-symbols must be functions of micro-weights and micro-probabilities that preserve relevance ordering and do not flip under benign regrouping. This is not a requirement that weights remain numerically identical under aggregation; it is a requirement that the induced ordering and relative distinctions be stable under coarse-graining consistent with the same underlying dynamics.

M3: Stability under recursive reuse

Coherence is defined in terms of recursive reinforcement; what persists is what reappears as enabling structure across iterations. A weight function that swings wildly under small perturbations fails to track recursion; it tracks noise. Weights therefore must be tied to long-run contribution rather than momentary salience. This is a stability constraint. If a symbol's role in preserving compressive recoverability is consistent, its weight should not vary erratically with minor changes in sampling or transient context.

M4: Empirical derivability without exogenous semantics

There must exist at least one operational route by which *w* can be computed from observed data through a coherence proxy, without relying on interpretive judgments of meaning. This does not deny that meaning exists at higher levels of description; it denies that the theory may assume it to avoid hard inference. Without an empirical derivation route, coherence weighting remains unfalsifiable because any mismatch can be attributed to "wrong relevance assignments."

M5: Domain-translation invariance

Coherence is claimed to be cross-domain. If a domain mapping ϕ preserves coherence up to monotone transformation, then induced weights must be compatible under that mapping as well, again up to monotone transformation. This prevents a hidden collapse; relevance must not be smuggled in as domain-specific semantics that only looks universal because it is verbally described as such. A weighting that cannot survive translation is not coherence-derived; it is context-customized.

These commitments do not assert that w is uniquely determined by coherence. They assert something weaker and more defensible; w is *admissible* only if it satisfies these constraints. Within the admissible class, estimation may vary. Outside the admissible class, the framework ceases to be operational.

A central technical obstacle is that coherence, as previously defined, is a property of system configurations or state transitions, while coherence weighting in information theory attaches to source symbols or features. This mismatch is not a flaw; it is a translation problem. The correct bridge is not to pretend the mismatch does not exist, but to define a relevance functional that maps from state-level coherence dynamics to symbol-level contributions.

The basic intuition is simple. A symbol is coherence-relevant to the extent that its presence supports recursive compressive recoverability of system trajectories, and coherence-irrelevant to the extent that its removal leaves coherence unchanged. This can be made operational by measuring the marginal contribution of a symbol or feature to a coherence proxy computed over time.

A derivation pipeline: $data \rightarrow \hat{C} \rightarrow \rho \rightarrow w$

The following pipeline is intentionally generic. It is designed to be instantiated across substrates while preserving the same inferential logic.

Step A: Select a state stream.

Choose a time-indexed stream X_t representing the system's configuration at each step. The representation may be discrete symbols, feature vectors, graphs, or structured logs,

provided that it captures the system's evolving configuration and it is stable enough that compression-based estimators are meaningful. The pipeline does not require that X_t be exhaustive; it requires that X_t be consistent.

Step B: Estimate proxy coherence via compression deltas.

Let $K(\cdot)$ be a compression-based complexity estimator. The choice of K is deferred; what matters initially is that it is consistent and applied uniformly across the stream.

Define a per-step compression delta,

$$\Delta K_t = K(X_t) - K(X_{t+1}),$$

and a bounded proxy coherence,

$$\hat{C}_t = \sigma\left(\alpha \cdot \frac{\Delta K_t}{K(X_t) + \varepsilon}\right),$$

where σ maps to $[0,1]$, α is a scaling constant, and ε stabilizes the denominator. This construction encodes a minimal claim; coherence increases when the system's evolving configurations become more recursively compressible and decreases when they fragment into less recoverable structure. The proxy is not intended to be metaphysically exact. It is intended to be operationally stable.

Step C: Define symbol-level relevance by ablation contribution.

Coherence weighting requires a relevance score $\rho(x)$ for each symbol or feature x. A non-semantic, structural route is leave-one-out contribution. Construct an ablated stream $X_t^{(-x)}$ in which x is masked, removed, or neutralized, using a method appropriate to the representation (masking tokens, zeroing features, removing edges, etc.). Compute the corresponding proxy coherence $\hat{C}_t^{(-x)}$ using the same estimator K and normalization. Define

$$\rho(x) = \mathbb{E}_t\left[\hat{C}_t - \hat{C}_t^{(-x)}\right].$$

If removing x reduces proxy coherence on average, then x contributes positively to recursive recoverability and is coherence-relevant. If removing x has negligible effect, x is coherence-irrelevant. If removing x increases coherence, x functions as coherence-negative noise

relative to the current dynamics. This definition does not require that x "mean" anything. It requires only that x participates in structural reinforcement detectable through compression-based recoverability.

Step D: Induce bounded weights from relevance.

Let μ be a monotone bounded mapping, and define

$$w(x) = \mu(\rho(x)).$$

A concrete admissible choice is

$$w(x) = \sigma(\beta\rho(x)),$$

where σ bounds to $[0,1]$ and β sets sensitivity. This step enforces M1 explicitly. The resulting weights are derived, not assigned. They are forced by measured marginal contribution to proxy coherence under an explicit ablation procedure.

Step E: Compute coherence-weighted information quantities and compare.

With induced weights $w(x)$, coherence-weighted entropy, mutual information, and channel capacity become computable from data. The prior formal results then apply; coherence-weighted channel capacity and selective compression bounds are not philosophical gestures but experimentally estimable quantities. One may now compare coherence-weighted coding schemes to Shannon-optimal coding schemes under controlled conditions, including distribution shift and constraint perturbation.

This pipeline accomplishes the core objective; it closes the gap between the formal machinery of coherence-weighted information theory and the empirical status of the weights that drive it.

Falsifiable predictions produced by induced weighting

The pipeline is not merely constructive; it is testable. The most important consequence of making w induced rather than assigned is that the theory becomes vulnerable in the right way. The following predictions are not rhetorical; each can fail.

P1: Coherence-weighted compression preserves coherence-bearing structure better at equal bitrate under shift.

If w is induced from recursive coherence contributions, then compression optimized for coherence-weighted entropy should preserve the system's coherence-bearing structure more reliably than Shannon-optimal compression at the same bitrate, particularly under distribution shift where purely frequency-based coding preserves high-probability noise and discards low-frequency structural signals. This prediction is falsified if coherence-weighted coding yields no systematic advantage, or if it consistently underperforms across domains and shifts.

P2: Weight rankings are stable under benign coarse-graining.

Compute weights at two resolutions: micro-symbols and macro-features (or grouped symbols). If the scale-consistency commitments are correct, relevance rankings should be approximately preserved under coarse-graining, up to monotone transformation. This prediction fails if relevance ordering flips unpredictably under regroupings that do not materially change the underlying dynamics.

P3: Declines in $\hat{C}$ precede regime transitions more reliably than standard metrics.

If coherence proxies track the approach to critical transitions, then declines in $\hat{C}$ should lead observed regime shifts, e.g. clinical episode onset, institutional breakdown, or ecological collapse, earlier than conventional symptom inventories or aggregate performance metrics. This prediction fails if $\hat{C}$ provides no lead-lag advantage or behaves as a lagging indicator.

P4: Constraint-relief interventions measurably increase coherence proxies relative to purely interpretive interventions.

If coherence failure is driven by constraint topology and option-space contraction, then interventions that materially alter constraints, such as reduced punishment density, shortened feedback loops, or expanding viable actions, should increase $\hat{C}$ and reshape induced weights in predictable ways. Interventions that alter narrative interpretation without

altering constraints should have weaker or less stable effects. This prediction fails if interpretive change alone reliably increases coherence proxies while structural constraint relief does not.

These predictions place the framework in contact with empirical refutation. The aim is not to guarantee success but to guarantee testability.

The residual hard seam: the choice of compression estimator K

Even with induced weighting, one seam remains, and it is the seam that distinguishes a universal invariant from an estimator-dependent artifact, the choice of compression estimator K. Different compression schemes privilege different regularities. A compressor that is good at capturing repetition may miss higher-order grammar; one tuned for local patterns may miss long-range dependencies. If coherence depends heavily on the idiosyncrasies of K, then the induced weights w inherit that idiosyncrasy, and the claim of cross-domain invariance becomes vulnerable.

The correct response is not to deny this seam. It is to reframe invariance as robustness. A coherence measure intended to be operational must not be invariant under all possible estimators, an impossible requirement, but it should be stable across reasonable estimator classes, at least up to monotone transformation. Concretely, if multiple compressors $K_1, K_2, \ldots$ yield coherence proxies whose trajectories disagree wildly on the same data, then coherence is not being estimated as a system property but as an estimator artifact. Conversely, if distinct estimator classes converge on similar coherence dynamics and similar induced weight rankings, the invariance claim gains empirical support.

This is not a philosophical escape hatch. It is an empirical program. The theory commits itself to a robustness criterion; coherence and induced weights must be estimator-stable in the regimes where the theory claims universality.

Conclusion

Coherence-weighted information theory becomes operational only when its weights are disciplined. The minimal commitments articulated here do not remove all estimation choices, but they eliminate arbitrariness by constraining what qualifies as coherence-derived. The derivation pipeline provides an explicit route from observed streams to coherence proxies, from coherence proxies to symbol-level relevance, and from relevance to induced bounded weights. The resulting framework is testable; it predicts measurable advantages for coherence-weighted compression under shift, stability of relevance rankings across scales, coherence declines as leading indicators of transitions, and differential impacts of constraint-relief interventions. The remaining seam, the compression estimator, does not invalidate the approach, but it forces a precise claim. Coherence is a robust system property only if it is stable across reasonable estimator classes. With that clarification, coherence weights cease to be a discretionary semantic overlay and become what the program requires them to be, a derived functional grounded in recursive structural reinforcement, fit for empirical convergence and falsification.

References

James, B. (2024). Achieving Coherence: Modeling Complexity in Dynamic Systems. Kindle Direct. Audible. https://philpapers.org/rec/JAMACM-3

James, B. (2025). Beyond Shannon: Coherence Information Theory and the Future of Communication. PhilPapers. https://philpapers.org/rec/JAMBSC

James, B. (2025). The Coherence Engine. PhilPapers. https://philpapers.org/rec/JAMTCE

James, B. (2025). Formal Foundations of Adaptive Coherence: A Recursive Metric of Reality. PhilPapers. https://philpapers.org/rec/JAMFFO

James, B. (2025). Formal Foundations of Coherence Information Theory: Capacity and Compression Theorems. PhilPapers. https://philpapers.org/rec/JAMFFO-2